

Amendments to the Specification.

Change the title of the patent application to

**SIMULTANEOUS MULTI-USER REAL-TIME VOICESPEECH  
RECOGNITION SYSTEM**

Substitute the following paragraphs of the specification:

[0003] This writing explains a method to solve the problems in generating a Multi User Conversational Voice Log or (MVL). In this document, the term speech recognition shall be used to denote conversion of spoken words to text. The term voice recognition will be used to denote identifying a person from the person's voice. Otherwise, however, the term voice shall be used to denote the actual sound a person makes when speaking or an accurate recording of that sound, unless another meaning is clear from the context. There are many problems and sub problems that need to be solved in order to create a MVL. These include:

Real-time voicespeech recognition and capture of many people

Distinguishing each person in a group individually

Creating the individual voice log

Integration of each person's voice log into a combined MVL

Organization of the many voice logs in the proper order

-Acceptable accuracy to make the log useful

-Making the text log easily accessible or printable on request

-Having a command set that can address the control of creating a MVL

[0007] It is a further object of the invention to provide a Simultaneous Multi-User

Real-time ~~Voice~~Speech Recognition System and text creator that is capable of being manufactured of high quality and at a low cost, and which is capable of providing a long and useful life with a minimum of maintenance. With these and other objects in view, as will be apparent to those skilled in the art, the invention resides in the combination of parts set forth in the specification and covered by the claims appended hereto, it being understood that changes in the precise embodiment of the invention herein disclosed may be made within the scope of what is claimed without departing from the spirit of the invention.

[0008] Using ~~voicesspeech~~speech recognition in the meeting environment combined with the ability to capture every person's conversation individually, including all people in total, allows minutes to be captured real-time and converted to the format of text. Thus creating a Multi-User Conversational Voice Log or "MVL". This concept can be used in many applications spanning from a single person's conversation to a meeting of the United States House of Representatives, and everything in between. Other features can be added to such a device, for example, real time language translation by displaying text in an alternate language from the input language. However, industry and people in general cannot take full advantage of ~~voicesspeech~~speech recognition because of many problems that exist with the existing technology.

[0009] The following sections will discuss the Conference To Text System (CTTS) , which are the hardware and software components that enable the ability to generate a Multi-user Voice Log or MVL. The invention described below addresses the

following problems:

Components of the technology that do not exist.

Existing components and technology have not been brought together and debugged to support this aspect of voicespeech recognition.

Training the many systems needed to recognize each person is time consuming and not feasible.

Lack of a command set to control creation of a MVL.

[0011] 1) Computer hardware with high performance that can service a person individually and collaborate in a high performance local area network environment. The hardware needs to have the power and packaging to be customer accepted. A unit containing a high-speed processor, memory, mass storage, audio input, optional display, and mouse would be used for each individual to be captured. A separate computer system 'functioning as a Voice Log Integrator is connected to the user units by a network (FIG. 3). An operating system and voicespeech recognition application is used on each unit. Voice Model Mobility allows users to obviate the need for training.

[0022] FIG. 6 shows the same internal components as FIG. 5 above, configured as a prototype Micro-Computer for handheld large vocabulary voicespeech recognition packaged in a form factor of standard PC 5 1/2 inch slot. It is shown with the display screen that can be located at a user location in a conference room. Below the pen is the microphone connector and enable switch. It also has a network port for

connecting back to the CTTS integrator. It can be used as touch screen, or with a keyboard and/or a mouse.

[0023] Computer hardware to support these types of applications must include at least the following components to be effective:

High-speed microprocessors with robust floating point features

Large on chip and/or off chip cache

High-capacity/fast main memory

Quality sound input device with performance focused in the range of the human voice or signal to be translated

-An operating system specifically configured (tuned) for the application of ~~voice~~speech recognition and data base management.

[0026] The Dialog Integrator is software that executes on the CTTS system (See F1G.1). It organizes the captured voice text or voice text logs and puts them in chronological order for screen output or export to a file. In summary, there are three items to be discussed with the Dialog Integrator. These three items include 1) Time stamp, 2) Integrating many voice logs together into a Multi-User Conversation Voice Log, and 3) Taking the voice text and index for each word/sentence and putting that into a database table, text file, or some other file/format. The log file contains a time stamp or some other method to synchronize all voice logs intended for conversion to MVL. The time stamp can be done as an integrator component, or the time stamp

may be placed into the log by the voicespeech recognition software or a related utility.

[0027] To date there are no voicespeech recognition software packages on the market that include indexing, or time stamping as the words are spoken or after a delay or pause of some amount of time, for example. The integrator could be built into other components like SVM or VMM as well.

[0031] Methods of indexing:

Counter

Clock

-Text character sequence

Control code sequence

4) Real-time voicespeech recognition captures software and components that maintain a reliable level of accuracy.

[0032] A real-time voicespeech recognition software package is needed to capture the voices. There are many voicespeech recognition packages on the market and in the public domain. The voicespeech recognition software must provide a consistent standard level of accuracy. It must also contain a very large vocabulary. The voicespeech recognition engine may be the best place to stamp the time or index because it is closest to the source. Another important feature is a consistent and reliable level of accuracy.

This is essential as the public becomes aware of how to speak to machines that recognize voice. This would allow the public to grow with the technology. The key component that could help this to work would be the use of a "Voice Accuracy Meter".

[0036] Key commands for a conference ~~voices~~speech recognition system:

Start meeting

Stop meeting

Recognize group

Recognize user

Pause meeting

Print meeting

Print individual "name"

Index method "type"

Strike class comment

Start mute

Stop mute

Start recognize "name"

Stop recognize "name"

Off the record

On the record

Bookmark "phrase"

Mark for correction

List uncorrected

List corrected

Play voice reference

Display user "name"

Applications

[0040] Normally, for ~~voices~~speech recognition software to support large vocabularies (30,000 plus words), training the software to recognize a person's voice accurately and consistently is ongoing because of ever-changing parameters with regards to the human voice and environment. Therefore, if the hardware and software (machine) that provides recognition is not current with the parameters of the person speaking, there is a delta between the user and the machine. This delta can be the cause of and a measure of inaccuracy. As people use different machines, the delta becomes dynamic and accuracy becomes inconsistent. To allow any user the ability to unplug the "Voice Model" and plug it into the current system in use allows the user the ability to have consistent accuracy. This concept was defined in a previous patent application by Darrell Poirier, and is labeled as Voice Model Mobility (VMM).

[0043] These examples explain in overview how a Super Voice Model could be designed or implemented, the concept being that many people use ~~voices~~speech recognition machines with no pre-training.

Replace the ABSTRACT with the following: